

Estimating Preferential Flow Paths within an Aquifer System

Margaret Short, Statistical Sciences Group, Los Alamos National Laboratory

Laura Guadagnini, Dipartimento di Ingegneria Idraulica, Ambientale Infrastrutture Viarie e Rilevamento (DIIAR), Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy

Alberto Guadagnini, Dipartimento di Ingegneria Idraulica, Ambientale Infrastrutture Viarie e Rilevamento (DIIAR), Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy

Daniel M. Tartakovsky, Department of Mechanical and Aerospace Engineering, University of California, San Diego, and Los Alamos National Laboratory

Dave Higdon, Statistical Sciences Group, Los Alamos National Laboratory

LA-UR-06-5233

Abstract

The subsurface environment beneath the Municipality of Bologna, Italy is comprised of a series of alluvial deposits which constitute large and productive aquifer systems. These are separated from the shallow, free surface aquifer by a low permeability barrier called aquitard Alpha. The upper aquifer contains water that shows relevant contamination from industrial pollutants. The deep aquifers are relatively pristine and provide about 80% of all groundwater used for drinking and industrial purposes in the area of Bologna. Hence, it is imperative that planners understand where along aquitard Alpha there exists potential direct connection between the upper and the deep aquifers, which could lead to contamination of the city's key water supply well fields.

In order to better understand the existence of preferential flow paths between these aquifer systems, we carry out a statistical analysis in which the aquitard is represented as a bivariate spatial process which simultaneously models its thickness as well as its sedimentological properties. The sedimentological component of the process provides a spatial map of the probability of coarse material regions within aquitard Alpha, which can induce direct vertical connection between the aquifers. This map is then cross referenced with other forms of data regarding the hydrology of the region.

1 Introduction

The key geological units beneath the Municipality of Bologna, Italy host large and productive aquifers, which are separated vertically from the upper, free surface aquifer by a low permeability barrier called aquitard Alpha (Figure 1). The upper aquifer contains water that shows evidence of contamination by industrial pollutants. The lower aquifers are relatively pristine and provide about 80% of all groundwater used for drinking and industrial purposes in the area of Bologna.

Relatively fast migration of contaminants between the upper and deep aquifers is possible through highly permeable inclusions (e.g., fractures, cracks, lenses of coarse material, etc.) embedded in aquitard Alpha. The ability to ascertain the locations and spatial extent of such inclusions which constitute preferential flow paths for contaminants is paramount for ensuring the safe and environmentally sound exploitation of groundwater resources in the region.

To identify highly permeable inclusions in the otherwise impermeable aquitard Alpha, we carry out a statistical analysis in which the aquitard and its hydrogeological attributes are treated as spatially correlated random fields. These fields are estimated using sedimentological and stratigraphic data. This analysis allows for the construction of a spatial map of the probability of regions consisting of coarse materials dominated by high hydraulic conductivities within the aquitard Alpha. This map is then cross referenced with other types of hydrogeologic information.

1.1 Description of the site and motivation of the study

The city of Bologna lies on the alluvial plain of the Reno river, in the Emilia Romagna region of Northern Italy. This study focuses on a nearby area (Figure 1a) of about $50km^2$, which contains three major well fields whose combined yield accounts for about 80% of the municipality's groundwater supply. The most important environmental problem suffered by the municipality of Bologna is groundwater contamination. In the last decades, industrial development has produced many sources of pollutants. High concentrations of organohalogenetic compounds (mainly chlorinated solvents) were first detected within the urban area of Bologna in the 1980s. While the concentrations of trichloroethylene have been decreasing since the early 1980s, the concentrations of perchloroethylene have been rapidly increasing (the behavior of these compounds is mainly related to their use in industrial processes requiring halogenetic solvents). Increased concentration levels of nitrates were also detected in the groundwater in the eastern side of the city. Furthermore, traceable concentrations of iron and manganese were detected in the main drinking well fields.

The geologic structure of the region's subsurface has been the subject of numerous investigations

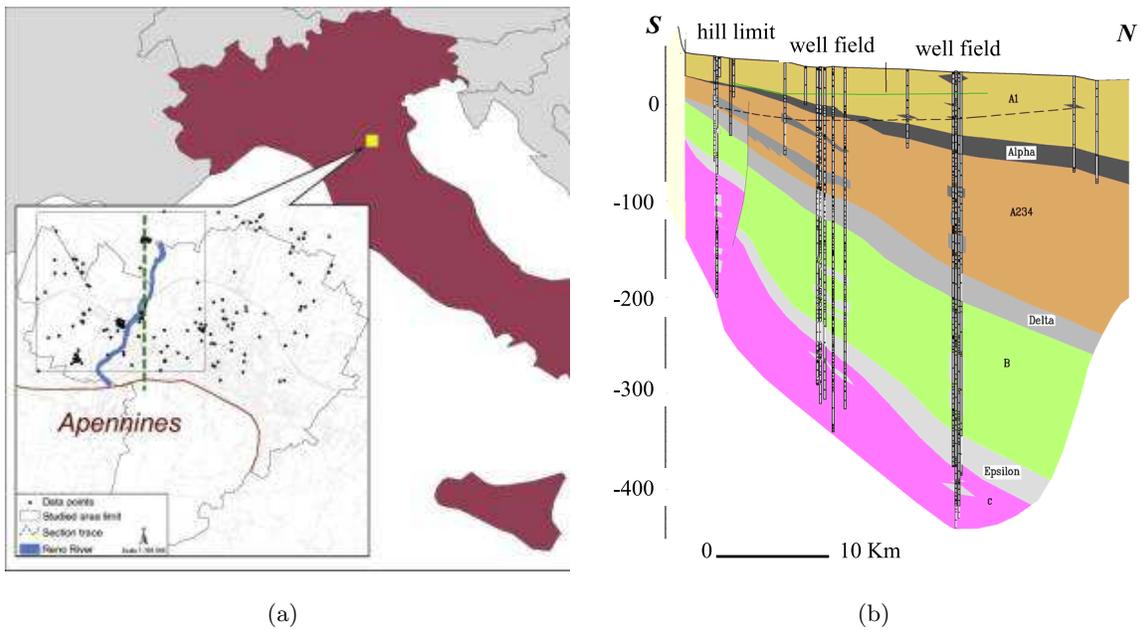


Figure 1: Locations of the municipality of Bolgna, the study area (the square region in the insert) and data points within the study area and its vicinity (a). A conceptual representation of the main geological units in the vertical cross-section of the study area (b). Our focus is on the analysis of the composition of an aquitard Alpha that separates two aquifers in the Bologna aquifer system.

(e.g., Ricci Lucchi, 1984; Ricci Lucchi et al., 1982; Francavilla et al., 1980; Amorosi and Farina, 1994a, 1994b, 1995). These and other studies reveal that the Reno river alluvial fan in the area has a wedge shape, whose thickness increases from South to North (Figure 1b). The alluvial fan rests on sea clayey deposits, which are rich with saline water. The water-bearing alluvial deposits in the Bologna area are more than 300m thick (Francavilla et al., 1980) and can be subdivided into three large-scale geological units (also known as depositional cycles) denoted by A, B, and C in Figure 1b. These depositional cycles, each about 100 – 150m thick, are separated by the clayey deposits denoted Delta and Epsilon, which form flow barriers and are commonly referred to as aquitards. The stratigraphic structure (spatial extent) of the major depositional units A, B, and C was investigated with a series of well logs (Regione Emilia Romagna, 1998). The same well log analysis revealed that the depositional cycle A can be further subdivided into two major subunits (denoted by A1 and A234 in Figure 1b), which are separated the aquitard called Alpha.

The alluvial deposits A234 and B form large and productive aquifer systems, which are heavily used by the municipality of Bologna as a major source of fresh water. Groundwater from the upper aquifer A1 shows signs of local contamination. Aquitard Alpha plays a major role as the natural barrier between the upper contaminated aquifer system A1 and the deeper aquifer systems A234 and B. Reliable identification of the spatial distribution of geologic facies (geomaterials) within this aquitard is crucial for identifying potential preferential paths for contaminant migration between the aquifers A1 and A234. It is also an essential component of the conceptual and mathematical models of subsurface water circulation in the region. Such an identification is hampered by the scarcity of available data. Probabilistic reconstruction of hydrofacies distribution within aquitard Alpha is the main objective of this study.

2 Sedimentological Data

A total of 39 logs of geognostical boreholes and 183 well logs were used to characterize the subsurface structure of the Reno river alluvial fan within the municipality of Bologna. Their locations are shown by the dots in Figure 1a. Of these, 123 well logs fall within the area under investigation — the region identified by the dashed-line square in the insert of Figure 1a — and are used in this analysis. From each of these well logs, two attributes of aquitard Alpha were extracted: its local thickness and the volumetric fraction (percentage) of the embedded coarse materials. These are reported as the cumulative thickness of fine materials (silt and clay) and the percentage occupied by coarse materials (sand and gravel), respectively. Figure 2 depicts spatial locations of the well

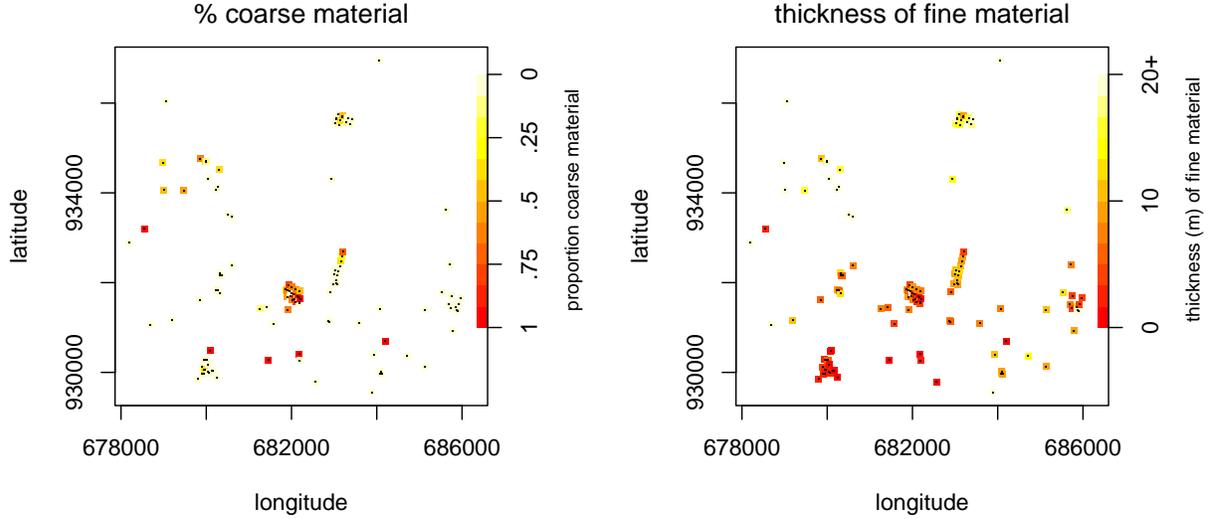


Figure 2: Spatial distribution of well core samples corresponding to aquitard Alpha. Left: the proportion of coarse material (sand and gravel) in the well core of aquitard Alpha. Right: the thickness of fine material (silt and clay) in the well core of aquitard Alpha.

core data over the study region.

The core data reveal that the thickness of aquitard Alpha varies from 1 – 3m in the areas near the peak of the Reno river alluvial fan (the southern region of the study area in Figure 1a) to 8 – 12m near the main municipal well-fields (the southwestern part of the study area in Figure 1a), to 30 – 45m in the Northern areas.

The data show that aquitard Alpha is composed mainly of fine (silty-clayey) material with local interbedding of coarse (sand and gravel) material. The dominance of the fine material defines the ability of aquitard Alpha to act as a natural barrier, which prevents flow and migration of contaminants between the upper aquifer and the deep groundwater reservoirs. The inclusions of coarse material are highly permeable and can effectively act as locations of preferential flow paths in the otherwise impermeable aquitard.

Aquitard Alpha is most likely to pose environmental problems (i.e., to display contaminant pathways) at locations where its relatively small thickness contains relatively large fractions of the coarse geomaterial. From each of the $n = 123$ well core samples the bivariate measurement

$$y(s_i) = \begin{pmatrix} y_T(s_i) \\ y_P(s_i) \end{pmatrix} = \begin{pmatrix} \text{thickness of fine material at } s_i \\ \% \text{ of coarse material at } s_i \end{pmatrix}, \quad i = 1, \dots, n \quad (1)$$

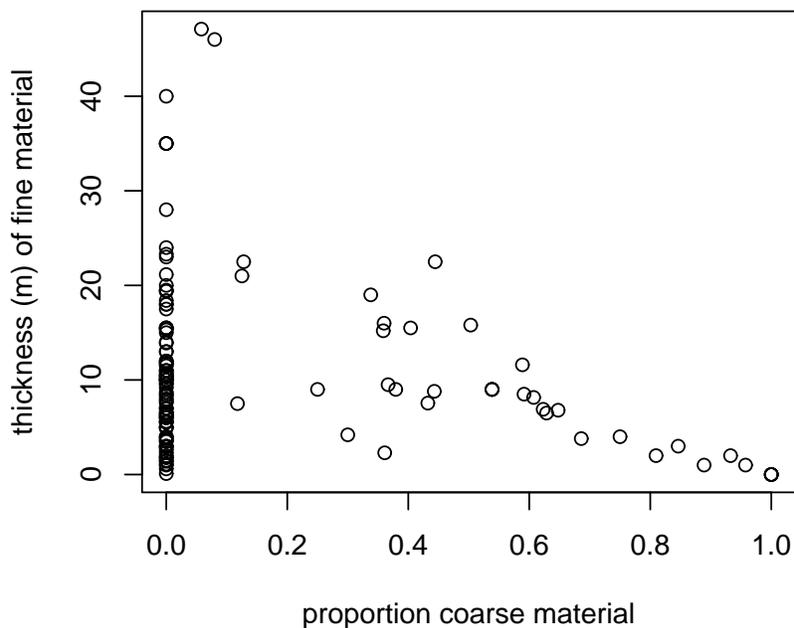


Figure 3: Proportion of coarse material (sand and gravel) vs. the thickness of impermeable material (silt and clay) in the well core samples of aquitard Alpha.

was recorded. The data are shown spatially in Figure 2 and as a scatterplot in Figure 3. The measurements showing aquitard Alpha to be locally thin as well as having a high proportion of coarse material suggest possible discontinuities within the impervious matrix of the aquitard. Hence sites for which the plotting symbols are dark for both plots of Figure 2 point to possible preferential flow paths between the upper and deep aquifers. Likewise, these same environmentally critical locations correspond to points in the lower right corner of Figure 3. In the next section we utilize these data, along with a binary spatial process model for the presence of preferential contaminant pathways through aquitard Alpha, to infer locations of possible discontinuities in the low conductivity structure of aquitard Alpha.

3 Statistical modeling

The statistical modeling for this application requires us to specify a spatial prior model for aquitard Alpha as well as a likelihood which determines how the observed data informs about the aquitard. Both of these model components are described in this section.

3.1 Spatial prior model for aquitard Alpha

Over the study region \mathcal{S} we focus on two features of aquitard Alpha: the total thickness of the fine materials; and the presence or absence of contaminant pathways through the aquitard. We specify a prior bivariate spatial process $(z_T(s), z_P(s))$, $s \in \mathcal{S}$, for aquitard Alpha. The first component, $z_T(s)$, which we call clay thickness, denotes the total thickness of impermeable material (clay and silt) within Alpha at spatial location s . The clay thickness process $z_T(s)$ is a positive, continuous field over the spatial region \mathcal{S} . The binary process $z_P(s)$ we call connectivity. It denotes the presence ($z_P(s) = 1$) or absence ($z_P(s) = 0$) of preferential contaminant pathways through aquitard Alpha at spatial location s . Both of these fields are non-Gaussian, but are constructed using latent Gaussian process models.

The clay thickness is derived by taking the positive part of a standard Gaussian process $u_T(s)$ so that

$$z_T(s) = \begin{cases} u_T(s) & \text{if } u_T(s) \geq 0 \\ 0 & \text{if } u_T(s) < 0. \end{cases}$$

The underlying Gaussian process $u_T(s)$ has a mean function $\alpha_0 + \alpha_1 s_2$ that depends linearly on the north-south spatial coordinate s_2 . The random part of $u_T(s)$ is constructed using a discrete representation given in Higdon (2002) which we define as follows. Let \mathcal{S} be the spatial region defined by the study region in this application and let x_1^T, \dots, x_K^T be iid $N(0, \lambda_T^{-1})$ random variables (“knots”) associated with sites $w_1, \dots, w_K \in \mathcal{S}$. The knot locations are taken to be an equally spaced 25×25 array over the study region \mathcal{S} . The spatial process $u_T(s)$ for $s = (s_1, s_2) \in \mathcal{S}$ is constructed using the representation

$$u_T(s) = \alpha_0 + \alpha_1 s_2 + \sum_{j=1}^K x_j^T k(s - w_j; \sigma_T),$$

where we take $k(\cdot; \sigma_T)$ to be a circular, bivariate normal density with standard deviation σ_T . The precision parameter λ_T controls the precision of the x_j^T 's, which in turn controls the marginal variance of the $u_T(s)$ process. A $\Gamma(a_T = 1, b_T = .001)$ prior is specified for λ_T . The parameters controlling the mean function, (α_0, α_1) are given independent $U(0, 40)$ priors. Finally, the kernel width parameter σ_T controls the range of spatial dependence for the $u_T(s)$ process. We specify a $U(.04, 1)$ for σ_T after rescaling \mathcal{S} to the unit square $[0, 1] \times [0, 1]$. The lower bound of 0.04 is equal to the minimum spacings in the underlying knot locations. Allowing σ_T to take values lower than this will result in “dead” regions in $u_T(s)$ that move towards its mean value of $\alpha_0 + \alpha_1 s_2$. If smaller values of σ_T are required, then a finer grid of knot locations will be needed. For the application

here, the 25×25 grid is sufficient. Note that increasing the knot density has very little effect since λ_T can be correspondingly increased, leaving the induced distribution for the latent process $u_T(s)$ essentially unchanged.

This discrete representation allows the thickness field to be controlled by mean function parameters α_0 , α_1 , and the K knot values x^T . This is particularly useful for the MCMC approach used to explore the rather complicated posterior distribution resulting from this application (see (5)).

Conditional on the clay thickness field $z_T(s)$, the binary connectivity field $z_P(s)$ is constructed in a similar fashion using a latent 0 mean Gaussian process $u_P(s)$ according to the rule

$$z_P(s) = \begin{cases} 1 & \text{if } \beta_0 + \beta_1 z_T(s) + u_P(s) \geq 0 \\ 0 & \text{if } \beta_0 + \beta_1 z_T(s) + u_P(s) < 0. \end{cases} \quad (2)$$

Conditionally on $z_T(s)$, this is an example of a clipped Gaussian random field model from deOliveira (2000), where the mean function depends on clay thickness $z_T(s)$. It also can be considered a spatial generalization of the latent probit model (Agestri, 1990; Johnson and Albert, 1999). Unconditionally, this formulation results in a bivariate, non-Gaussian spatial model. The process $u_P(s)$ is defined in a manner analogous to that of $u_T(s)$:

$$u_P(s) = \sum_{j=1}^K x_j^P k(s - w_j; \sigma_P),$$

where the knot locations are the same as those for $u_T(s)$. Here the underlying knot values x_1^P, \dots, x_K^P are given independent $N(0, \lambda_P^{-1})$ distributions a priori. The precision parameter λ_P is fixed so that the resulting Gaussian process $u_P(s)$ has a nearly constant marginal variance of 1. This variance is not exactly constant over \mathcal{S} due to the discrete representation. However, the fluctuations over \mathcal{S} are small enough to be negligible. Also, the kernel $k(\cdot, \sigma_P)$ is controlled by its own scaling parameter σ_P . As with its clay thickness counterpart, σ_P is assigned a $U(.04, 1)$ prior.

In equation (2), β_0 controls the mean connectivity probability; β_1 controls the dependence between the clay thickness and connectivity fields. We expect that β_1 will be negative, so that greater clay thickness at spatial location s makes it less likely that there is a preferential pathway through aquitard Alpha (i.e., $z_P(s) = 1$). However, we specify wide $U(-15, 15)$ priors for both β_0 and β_1 to allow the well core data to inform about these parameters.

To visualize how this model works, Figure 4 shows the grids used to create the two processes, $u_P(s)$ and $u_T(s)$, along with realizations for these two processes. The final column of this figure shows the induced realization for the two dependent fields, $z_T(s)$ and $z_P(s)$. Note that the realization of $z_T(s)$ has been truncated to be strictly positive, and the realization from $z_P(s)$ is binary,

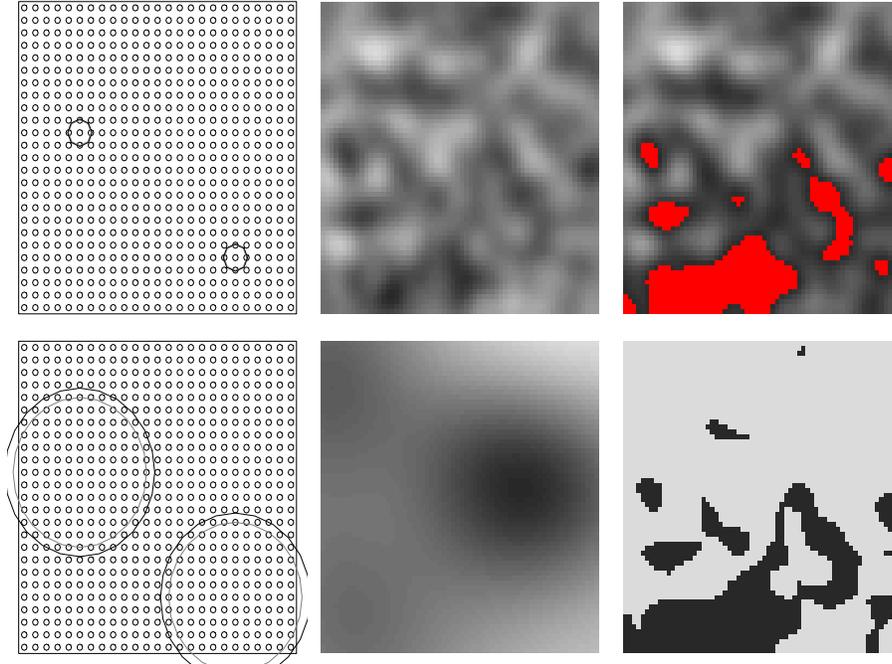


Figure 4: Construction of the bivariate spatial model. Independent normal variates are located at the grid locations shown in the left hand column of plots. These “knot” values are convolved by a normal kernel whose standard deviation is given by the circles in the plots of the first column. The induced fields $u_T(s)$ and $u_P(s)$ are shown in the middle columns. The right hand columns show the two spatial fields $z_T(s)$ and the binary connectivity field $z_P(s)$. The red portion of the top right image shows where $z_T(s) = 0$.

constructed from $z_T(s)$ and $u_P(s)$. For this realization, values for the parameters x^T , x^P , σ_T , σ_P , β_0 , and β_1 were taken from a posterior realization used in the eventual analysis of aquitard Alpha.

3.2 Incorporating well core data

The data inform about the underlying processes through the likelihood. For this application, the likelihood factors into two terms: one involving the clay thickness, and one involving the connectivity field. The first term is a standard likelihood for a spatial model where the observed clay depth y_{i1} is a noisy version of the true depth. The observed clay depth is not expected to be exact since the well core data typically carries a fair bit of uncertainty. In addition, we expect this noise term to absorb some of the small scale variability in the actual clay depth. Taking $y_T =$

$(y_T(s_1), \dots, y_T(s_n))^T$ to be the n -vector of clay depth observations and $z_T = (z_T(s_1), \dots, z_T(s_n))$ to be the clay depth process at the observation sites, this component of the likelihood becomes

$$L_T(y_T|z_T, \lambda_y) \propto \lambda_y^{\frac{n}{2}} \exp \left\{ -\frac{\lambda_y}{2} (y_T - z_T)^T (y_T - z_T) \right\}$$

where the parameter λ_y controls the observation precision.

The second component of the likelihood determines how the observed data inform about the binary connectivity process $z_P(s)$. Here it is standard to specify a sampling model for which observations $y(s_i)$, $i = 1, \dots, n$, are independent given the underlying spatial field so that

$$L_P(y(s)|z_P(s)) = \prod_{i=1}^n L_P(y(s_i)|z_P(s_i)). \quad (3)$$

Similar binary classification applications from spatial and image applications typically specify a normal sampling model where the mean of $y(s)$ depends on the state of $z_P(s)$ (Besag *et al.*, 1991; Hurn, 1998). However the nature of the bivariate measurements taken here, along with their spread – evident in Figure 3 – are incompatible with a normal sampling model for the $y(s_i)$'s.

If we assume the product form of the sampling model in (3), the fact that each $z_P(s_i)$ is binary means that the likelihood depends only on the ratios

$$L_P(y(s_i)|z_P(s_i) = 1)/L_P(y(s_i)|z_P(s_i) = 0) = r(y(s_i)), \quad i = 1, \dots, n$$

since (3) can be rewritten as

$$L_P(y(s)|z(s)) = \prod_{i=1}^n r(y(s_i))^{z_P(s_i)} L_P(y(s_i)|z_P(s_i) = 0) \propto \prod_{i=1}^n r(y(s_i))^{z_P(s_i)}. \quad (4)$$

Hence the data only inform about $z_P(s)$ through the specification of the *odds map* $r(y(s))$.

It remains to specify $r(y)$ as a function of possible bivariate outcomes $y(s) \in \{[0, \infty) \times [0, 1]\}$. In fact, we need only consider thicknesses between 0 and 50 meters since it is unlikely aquitard Alpha is thicker than 50 meters within the study region. In specifying the odds map, it is convenient to interpret $r(y(s))$ as the odds that the underlying $z_P(s) = 1$ at the spatial location s , independent of spatial information encoded into the prior for $z_P(s)$.

We rely on expert knowledge from hydro-geologists familiar with the study region to determine the odds map $r(y(s))$. We model $\log r(y(s))$ as a linear function of the observed clay thickness $y_T(s)$ and the observed proportion of coarse material in Alpha $y_P(s)$. The expert knowledge is used to specify a line for which the odds are equal to 1, and a slope which determines how quickly the odds change as the data move away from the $r(y(s)) = 1$ line. The best determination of this

function is given in the central frame of Figure 5. Considerations for determining this map include the following points.

- For the very small local thickness of aquitard Alpha ($< 10\text{m}$), even a very large fraction of fine material does not guarantee the spatial continuity of the aquitard.
- An intermediate thickness guarantees the continuity only for a relatively small fraction of coarse material. This is because coarse materials within a clearly identifiable aquitard manifest themselves as a sequence of interbedding structures, so that a porous pathway through the aquitard is possible.
- A very large fraction of coarse materials tends to be indicative of local discontinuities, unless the thickness of an aquitard is very large ($> 20\text{m}$).

In order to investigate sensitivity to this choice of odds map, 8 additional specifications are also given in the Figure 5 based on the expert knowledge. Each allows a perturbation of the $r(y(s)) = 1$ line and the slope.

3.3 Posterior distribution

After specifying common, independent gamma priors for the precisions and common uniform priors for the kernel width parameters σ_T and σ_P , the resulting posterior distribution has the form

$$\begin{aligned}
 \pi(\lambda_y, x_T, \lambda_T, \sigma_T, x_P, \sigma_P, \beta_0, \beta_1 | y) \propto & \quad (5) \\
 & \lambda_y^{\frac{n}{2}} \exp\left\{-\frac{\lambda_y}{2}(y_T - z_T)^T(y_T - z_T)\right\} \times \prod_{i=1}^n r(y(s_i))^{z_P(s_i)} \\
 & \times \lambda_y^{a_y-1} \exp\{-b_y \lambda_y\} \times \lambda_T^{\frac{K}{2}} \exp\{-\frac{1}{2}\lambda_T \|x_T\|^2\} \times \lambda_T^{a_T-1} \exp\{-b_T \lambda_T\} \\
 & \times \lambda_P^{\frac{K}{2}} \exp\{-\frac{1}{2}\lambda_P \|x_P\|^2\} \times \lambda_P^{a_P-1} \exp\{-b_P \lambda_P\} \\
 & \times I[.04 \leq \sigma_T, \sigma_P \leq 1] \times I[-15 \leq \beta_0, \beta_1 \leq 15]
 \end{aligned}$$

This distribution is sampled using Markov chain Monte Carlo with random walk Metropolis updates for the knot parameters x_T and x_P . Such updates are also used for the kernel parameters σ_T and σ_P , as well as β_0 and β_1 . Because the full conditionals for the precision parameters λ_y and λ_T are gamma, these parameters can be updated using Gibbs updates.

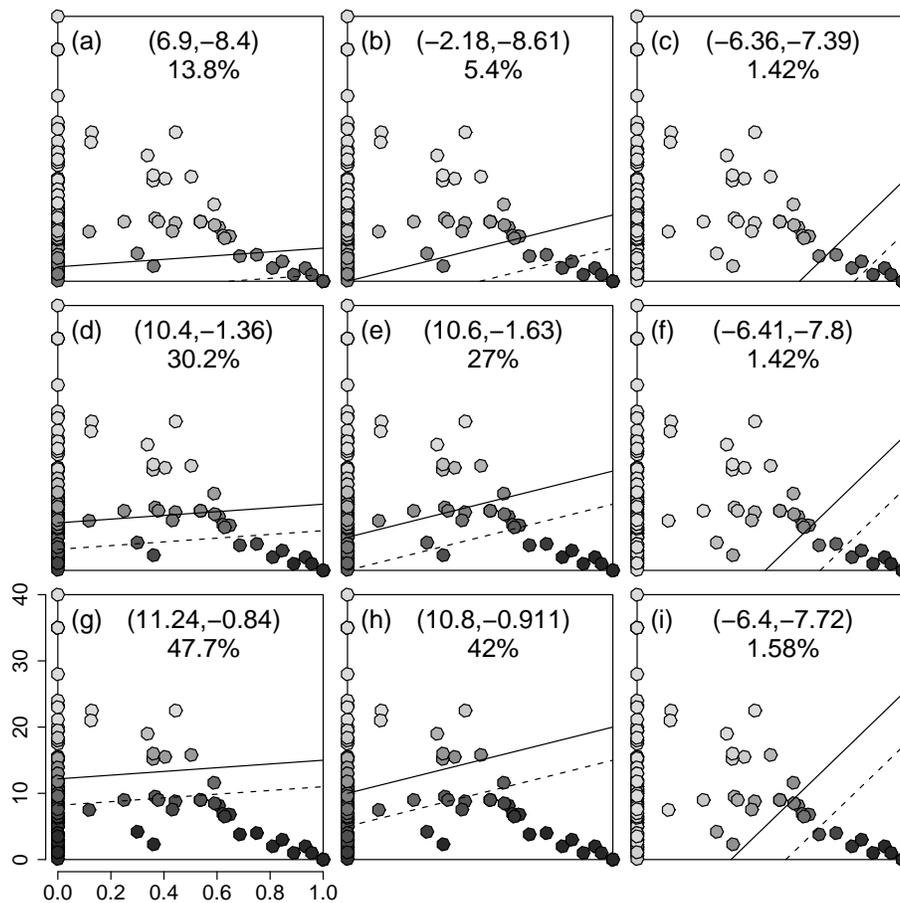


Figure 5: Odds maps based on expert judgement. The central figure corresponds to the best estimate; the remaining figures show other plausible odds maps. The solid line corresponds to data for which the odds of connectivity is 1. The dotted line corresponds to an odds of 10. For each odds map, the posterior proportion of the study region that is permeable is given. Above that are the posterior mean estimates for (β_0, β_1) .

4 Results

Figure 6 shows the posterior mean for the clay thickness of aquitard Alpha. In general it is thinner near the south edge of the study region. Inference regarding the clay thickness field $z_T(s)$ is almost completely insensitive to the choice of odds map since nearly all of the information regarding this field is contained in the thickness component of the likelihood $L_T(y_T(s)|z_T(s), \lambda_y)$.

The posterior mean for the binary connectivity field is given in Figure 7. Here the estimates vary considerably depending on which odds map is used. Also, given in Figure 5 are the posterior mean estimates for β_0 and β_1 as well as the total proportion of the study site where possible interconnections between upper and deep aquifers are revealed under each of the 9 different analyses.

The reconstructions of connected regions show an increased probability of discontinuity for the aquitard in the southern region of the study region. The magnitude of this probability varies depending on the odds map specified. Also, some of the analyses show potential for discontinuity in Alpha in the middle of the left hand side of the region. In Figure 7, the dark red regions denote probabilities greater than .9; the bright red regions denote probabilities greater than .95.

As expected, the posterior connectivity reconstruction depends a great deal on the assumed odds map which specifies how the data inform about the connectivity features of aquitard Alpha. Note that the reconstructions corresponding to the odds maps in the right hand column of Figure 5 show almost no risk of connectivity. This is because there are fewer data locations that give a high odds of connectivity. In addition, these odds maps lead to very little spatial continuity (see the data points in Figure 7), leading to spatially homogeneous reconstructions with a small posterior probability of connectivity. In essence, these rather extreme odds maps result in a constant posterior connectivity map which isn't consistent with the expert knowledge. Similarly, the overly strong odds maps given in Figures 5(g) and 5(h) lead to reconstructions that allow what is thought to be too large a region occupied by coarse geomaterials in aquitard Alpha.

The preferred odds map given in Figure 5(e) identifies potential connectivity in the southern part of the study region. The odds map corresponding to Figure 5(d) gives similar results. Both of these analyses point to a potential region of direct connection between the upper and deep aquifers in the central-western part of the study area.

The posterior distribution favors a rather short spatial correlation distance for the clay thickness field $z_T(s)$, and a large correlation distance for the process $u_P(s)$ used to build the connectivity field $z_P(s)$. A posterior realization of these fields are given in Figure 4. Residuals for the model fit are shown in Figure 8. For clay thickness, the residuals are simply $y_T(s_i) - \hat{z}_T(s_i)$, $i = 1, \dots, n$,

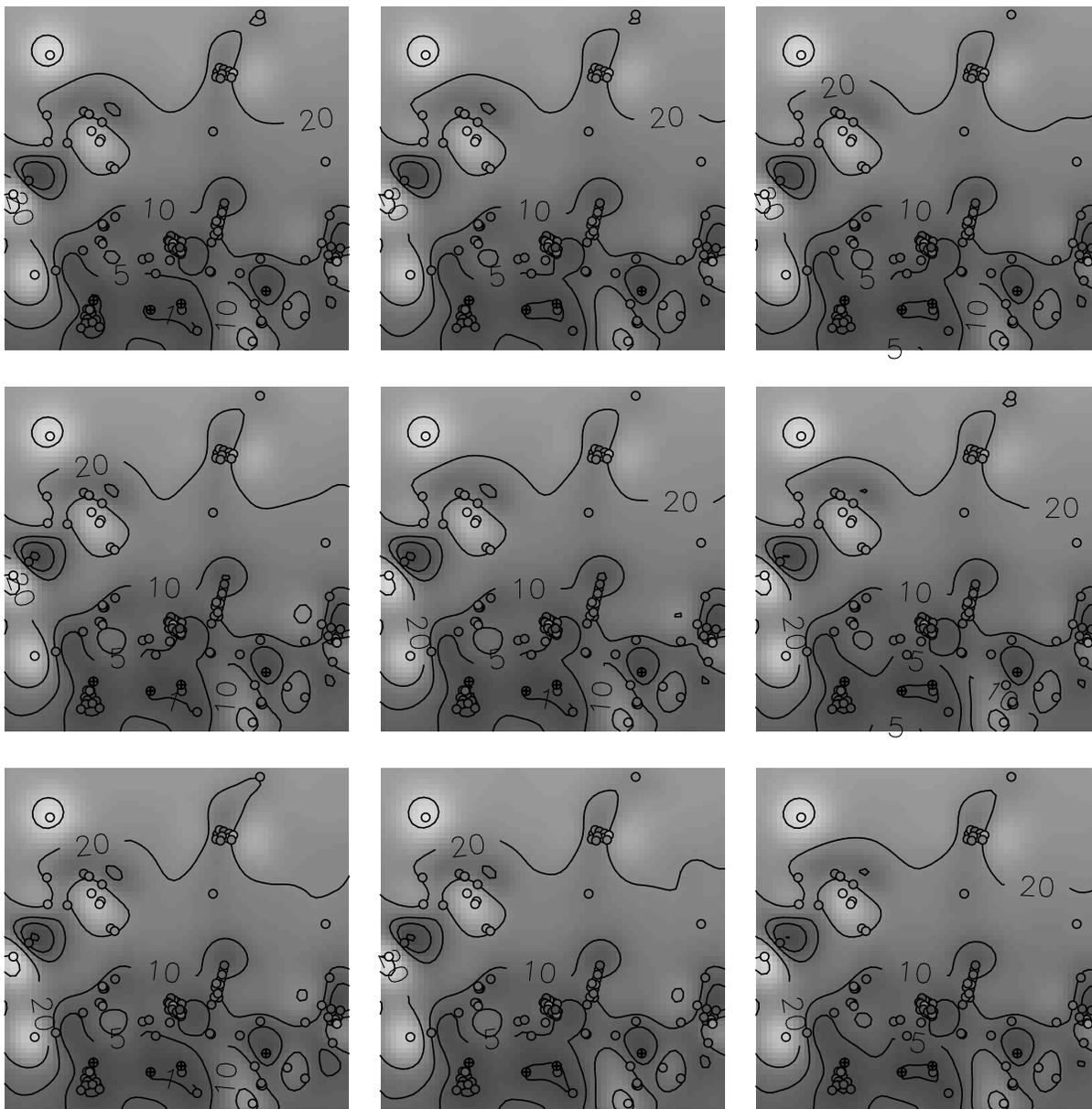


Figure 6: Posterior mean estimate for the clay thickness field $z_T(s)$. This estimate is not sensitive to the choice of odds map used from Figure 5.

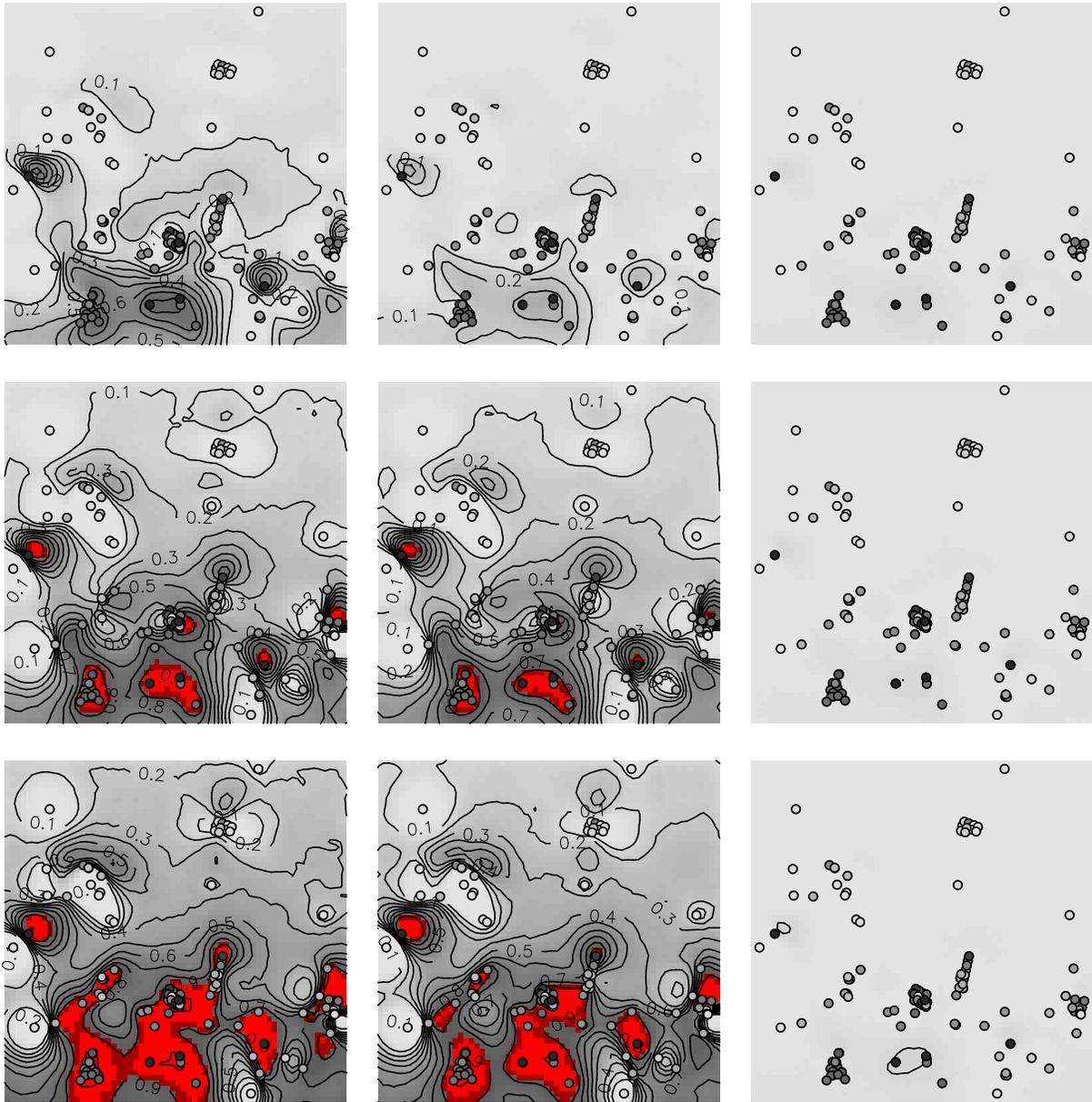


Figure 7: Pointwise posterior probability of connectivity for aquitard Alpha. The 9 posterior maps correspond to the posterior expectation of $z_P(s)$ using the corresponding odds maps given in Figure 5.

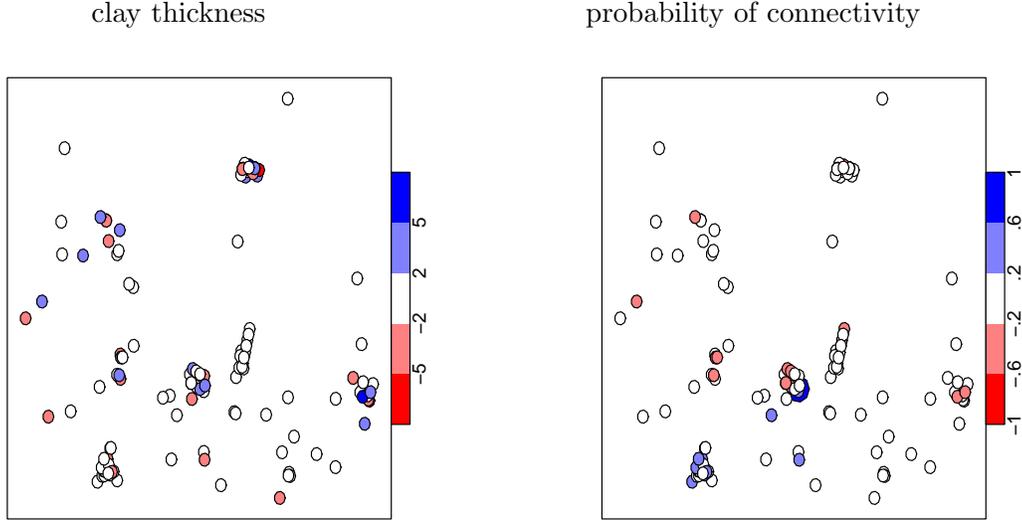


Figure 8: Residuals from the posterior mean for $z_T(s_i)$ and $z_P(s_i)$. For clay thickness, the residuals are simply $y_T(s_i) - \hat{z}_T(s_i)$, $i = 1, \dots, n$. For connectivity the residual is defined to be the difference between the inferred probability from the local well core that the aquitard is permeable at s_i , $p(s_i) = r(s_i)/(1 + r(s_i))$, and the posterior probability that $z_P(s_i) = 1$.

where $\hat{z}_T(s_i)$ denotes the posterior mean of the thickness field at s_i . For connectivity, the residual is defined to be the difference between the inferred probability from the local well core that the aquitard is permeable at s_i , $p(s_i) = r(s_i)/(1 + r(s_i))$, and the posterior probability that $z_P(s_i) = 1$. The posterior mean for these two fields appears to match the data well.

We used normal kernels in the convolution representation of both latent fields, $u_T(s)$ and $u_P(s)$, for computational reasons. The nonstandard model formulations of Section 3.1, as well as the need for predictions over a fine grid of spatial locations, require that $z_T(s)$ and $z_P(s)$ can be quickly produced given the model parameters. The convolution construction of the latent processes answers both of these needs. The resulting processes $u_T(s)$ and $u_P(s)$ are, to a close approximation, stationary zero mean Gaussian processes with a Gaussian covariogram. Such processes give very smooth realizations for $u_T(s)$ and $u_P(s)$ (Paciorek and Schervish, 2004). If one were to prefer rougher realizations for the processes, this would require a very dense set of knot locations, making the computations infeasible for this application given our current computing capabilities.

Given the high level of noise in the well-core data, and the fact that the connectivity field is not directly observed, we do not expect the data to inform about the smoothness of $z_P(s)$. Simulation studies have also lead us to conclude that the data do not inform about the covariance model for $z_T(s)$. However we can check the sensitivity of the clay thickness reconstruction to the choice

4.1.1 Comparison to piezometer readings

Since discontinuities in aquitard Alpha connect the upper and lower aquifers, water pressure (hydraulic head) in both aquifers is at equilibrium if there is a local discontinuity. This can be ascertained by piezometers which are capable of measuring hydraulic head in the upper and lower aquifers at (approximately) the same locations. Even though the regional network of Bologna comprises a large number of piezometers (Regione Emilia Romagna, 1998), only two pairs of piezometers satisfy this requirement. Their locations are denoted by the symbols 1 and 2 in Figure 9. We analyzed the average difference between hydraulic heads in the upper and lower aquifers Δh at these two locations over the period of 1999 – 2000. The first pair of piezometers (symbol 1, designated by 4030P and 4028P in the data base of the municipality of Bologna) gives $\Delta h = 27.53m$. The second pair (symbol 2, designated by 5261Pa and 5261Pb in the data base of the municipality of Bologna) gives $\Delta h = 35.37m$. Such large differences in hydraulic heads, combined with the remaining piezometric readings, strongly support the continuity of aquitard Alpha at these two locations. This is consistent with the central reconstruction of Figure 7 (as well as the others) which gives a posterior probability of .08 and .41 at those two locations. These data can also be incorporated into the data analysis by enforcing the condition $z_P(s) = 0$ at these two locations. The resulting connectivity estimate—computed via importance sampling using the original MCMC sample—is given in the right hand frame of Figure 9.

4.1.2 Comparison to qualitative geological cross-sections

The geological data set consists of the complete 123 well logs (stratigraphic columns), from which the sedimentological data described in Section 2 have been extracted. By supplementing this with the knowledge of the dynamics of depositional processes over the geological time scale, qualitative, expert-based reconstructions of geological cross-sections are produced. Six of these are shown in Figure 10. In these cross-sections, the dark color indicates the fine material inclusions into the coarse material depositional structures. The central figure shows the posterior mean connectivity field, along with the spatial locations of these cross-sections.

Cross-sections S1–S4 located in the southern part of the study area show little or no fine materials at the depths where aquitard Alpha should be. In contrast, the northern cross-sections S5 and S6 show very noticeable dark regions in the shallow depths where aquitard Alpha is located. Cross-section S1 identifies the presence of the fine materials at depths between 25 and 54m, which are significantly lower than the position of aquitard Alpha. Cross-section S2 shows highly localized

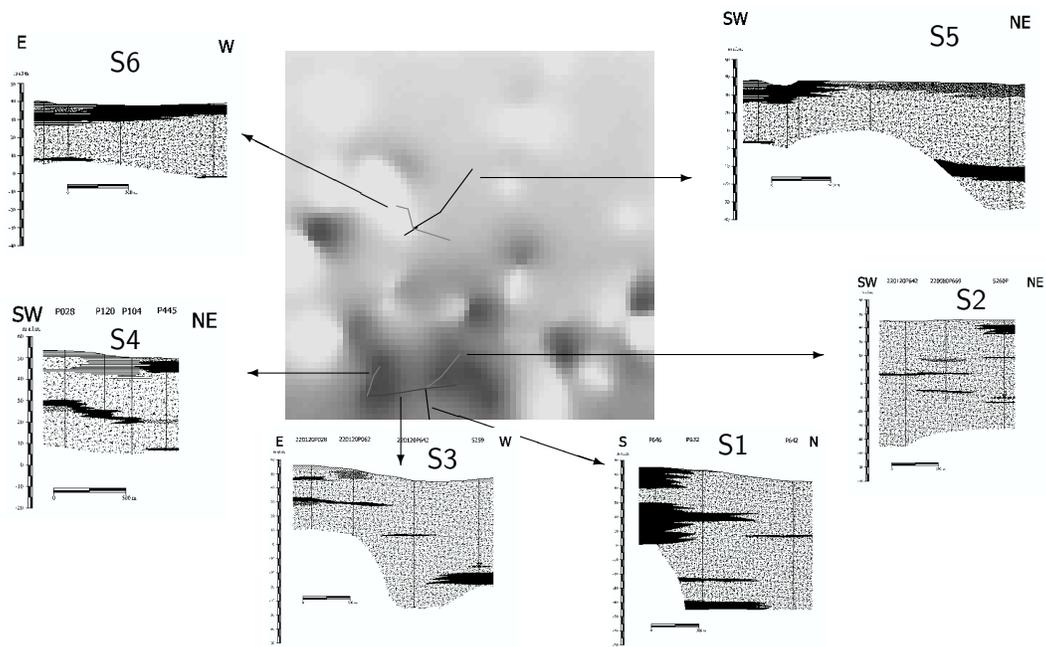


Figure 10: Qualitative geological sections. The posterior mean estimate of the connectivity field (center) along with qualitatively estimated cross-sections of the geology. These qualitative cross-sections were constructed using the 123 well cores described in Section 2 along with consideration of the dynamics of depositional processes over the geological time scale. Dark regions in the cross-sections denote fine-scale (low-conductivity) inclusions. In the central connectivity map, darker regions denote areas where aquitard Alpha is likely to be more permeable.

inclusions of fine materials, whose thickness is generally less than $1m$, into $90m$ of the ambient coarse material. Similarly, cross-section S3 reveals the small thickness and discontinuous nature of the fine material deposits within aquitard Alpha, mainly located at depths between 23 and $25.6m$. The thickness of the fine material deposits diminishes towards West. Near the Reno river (the western edge of cross-section S3), fine materials are not detectable at depths typical of aquitard Alpha. In cross-section S4, the fine material deposits belonging to aquitard Alpha are discerned at depths between 23 and $30.6m$. These clayey sediments tend to deepen when moving toward the NE edge, while at the same time, their thickness lessens, reaching about $10cm$ at the well log on the NE edge of the cross-section.

5 Discussion

This analysis has identified locations in the study region for which there is cause for concern regarding the ability of aquitard Alpha to protect the deep aquifers system from contaminants. Although the quantitative results of this analysis depend on the subjective odds map, the qualitative nature of the estimated permeable inclusions is fairly stable for 6 of the 9 odds maps. The southern region, along with some isolated regions of the study area show potential for connectivity between the upper and lower aquifers. The use of a more overtly subjective likelihood has proven useful in other scientific investigations as well; see Rappold *et al.*(2006), for example.

This study lead us to a new, non-Gaussian, bivariate spatial model that links a continuous, non-negative field with a binary field. Our original analysis considered only a clipped Gaussian field model for the binary connectivity field. However, this initial analysis proved unsatisfactory; it assigned appreciable chance of connectivity to locations in the northern region of the study area for which the clay depth is over $40m$. It was clear that a model that directly included dependence on the clay depth of aquitard Alpha was needed.

Acknowledgements

Part of the work, which includes stratigraphic and sedimentological data assimilation, was performed under the auspices of the European Commission (contract EVK1-CT-1999-00041-W-SAHaRA).

References

- Amorosi, A. and Farina, M. (1994a). Sequenze deposizionali nei depositi alluvionali quaternari del primo sottosuolo nell'area ad est di bologna, *Proceedings of the 1st European Congress on Regional Geological Cartography and Information Systems*, Vol. 5, pp. 35–54.
- Amorosi, A. and Farina, M. (1994b). Stratigrafia della successione quaternaria continentale della pianura bolognese mediante correlazione di dati di pozzo, *Proceedings of the 1st European Congress on Regional Geological Cartography and Information Systems*, Vol. 5, pp. 16–34.
- Amorosi, A. and Farina, M. (1995). Large-scale architecture of a thrust-related alluvial complex from subsurface data: the quaternary succession of the po basin in the bologna area (Northern Italy), *Giornale di Geologia* **57**: 3–16.
- Besag, J., Green, P. J., Higdon, D. M. and Mengersen, K. (1995). Bayesian computation and stochastic systems (with discussion), *Statistical Science* **10**: 3–66.
- Besag, J., York, J. and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion), *Annals of the Institute of Statistical Mathematics* **43**: 1–59.
- De Oliveira, V. (2000). Bayesian prediction of clipped Gaussian random fields, *Computational Statistics & Data Analysis* **34**(3): 299–314.
- Francavilla, F., D'Onofio, S. and Toni, G. (1980). Hydrogeological, structural, paleo-ecological and stratigraphical characteristics of the reno alluvial fan (bologna), *Quaderni Ist. Ric. Acque C.N.R.* **51**: 81–95.
- Guadagnini, L., Farina, M., Frontini, S. and Simoni, M. (2002). Geostatistical modeling of a heterogeneous alluvial aquifer, *Proceedings of the International Conference on Calibration and Reliability in Groundwater Modeling*, ModelCARE 2002, Prague, pp. 167–171.
- Hastie, T., Tibshirani, R. and Friedman, J. H. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York.
- Higdon, D. (2002). Space and space-time modeling using process convolutions, in C. Anderson, V. Barnett, P. C. Chatwin and A. H. El-Shaarawi (eds), *Quantitative Methods for Current Environmental Issues*, Springer Verlag, London, pp. 37–56.

- Hurn, M. (1998). Confocal fluorescence microscopy of leaf cells: An application of Bayesian image analysis, *Journal of the Royal Statistical Society, Series C: Applied Statistics* **47**: 361–377.
- Johnson, V. E. and Albert, J. H. (1999). *Ordinal Data Modeling*, Springer-Verlag Inc.
- Paciorek, C. J. and Schervish, M. J. (2004). Nonstationary covariance functions for gaussian process regression, in S. Thrun, L. Saul and B. Schölkopf (eds), *Advances in Neural Information Processing Systems 16*, MIT Press, Cambridge, MA.
- Rappold, A. G., Lavine, M. A. and Lozier, S. (2006). An assessment of climate change in the ocean, mixed-layer depth, and subjective likelihood, *Journal of the American Statistical Association* to appear.
- Regione Emilia Romagna (1998). Riserve idriche sotterranee della regione emilia romagna (Groundwater resources of the Emilia Romagna region), *Technical report*, Firenze.
- Ricci Lucchi, F. (1984). Flysh, molassa, clastic deposits: traditional and innovative approaches to the analysis of north Apeninic basins, *Cento Anni di Geologia Italiana*, Giubilare 1° Centenario, Societa Geologica Italiana, pp. 279–295.
- Ricci Lucchi, F., Colalongo, M. L., Cremonini, G., Gasperi, G., Iaccarino, S., Papani, G., Raffi, S. and Rio, D. (1982). Sedimentary and palaeogeographic evolution of the apenninic margin, *Guida alla geologia del margine appenninico padano*, Guide geologiche regionali, Societa Geologica Italiana, pp. 17–46.
- Winter, C. L., Tartakovsky, D. M. and Guadagnini, A. (2003). Moment equations for flow in highly heterogeneous porous media, *Surveys in Geophysics* **24**: 81–106.